学校编码：10384　　　　分类号 ______ 密级 ______

学号：200423024　　　　UDC______

厦门大学

硕　士　学　位　论　文

矩阵和与差的特征值及奇异值不等式

The Inequalities of the Singular Values and Eigenvalues of Summation and Difference of Matrices

林　智　鹏

指导教师姓名: 卢 琳 璋 教授

专 业 名 称: 计 算 数 学

论文提交日期: 2007 年 4 月

论文答辩日期: 2007 年　 月

学位授予日期: 2007 年　 月

答辩委员会主席: ____________

评　　阅　　人: ____________

2007 **年** 4 **月**

厦门大学学位论文原创性声明

兹呈交的学位论文，是本人在导师指导下独立完成的研究成果。本人在论文写作中参考的其它个人或集体的研究成果，均在文中以明确方式标明。本人依法享有和承担由此论文而产生的权利和责任。

声明人（签名）：

年　　月　　日

厦门大学学位论文著作权使用声明

本人完全了解厦门大学有关保留、使用学位论文的规定。厦门大学有权保留并向国家主管部门或其指定机构送交论文的纸质版和电子版，有权将学位论文用于非赢利目的的少量复制并允许论文进入学校图书馆被查阅，有权将学位论文的内容编入有关数据库进行检索，有权将学位论文的标题和摘要汇编出版。保密的学位论文在解密后适用本规定。

本学位论文属于

1、保密（　　），在　　　年解密后适用本授权书。

2、不保密（　　）。

（请在以上相应括号内打"√"）

作者签名：　　　　　　　　日期：　　　年　月　日

导师签名：　　　　　　　　日期：　　　年　月　日

摘 要

矩阵特征值、奇异值的不等式的研究是矩阵论及矩阵扰动分析的主要课题之一. 它研究的问题一般包括单个矩阵的特征值及奇异值之间的关系，多个矩阵之间的特征值奇异值及它们界的估计问题等. 在早期， Wely,Fisher,Wielandt 和 Hoffaman 等人在此领域做了大量的工作. 1953 年， A. J. Hoffaman 和 H.W.Wielandt 给出了两个对称矩阵差的特征值不等式，即著名的 Wielandt-Hoffaman 定理. 它将特征值的摄动与摄动矩阵的 Enclid 范数联系起来, 是对以浮点算术运算的正交变换为基础的误差分析是最有用的结果 [15].

本文是在 Wielandt-Hoffaman 定理的基础上，给出了它的反向不等式，作出了比较简洁的证明, 然后给出了两个矩阵和的特征值关系, 它的结论与 H-W 定理的形式几乎一样完美, 最后把它推广到奇异值的情形, 通过构造对称阵的方法巧妙的证明了奇异值不等式, 同样得到了两个矩阵和与差的奇异值不等式, 它的形式与 W-H 定理也是一致的.

本文可以分以下四个部分，第一节主要介绍相关的问题背景，并概述文章的主要内容. 在第二节里，引进了文章中要用到的一些定义，引理及基本定理. 第三节给出了 W-H 定理的对称形式并推广到矩阵和的形式. 第四节得到了任意 $m \times n$ 阶实矩阵和与差的奇异值不等式.

关键词：特征值；奇异值；不等式.

Abstrct

The studying of the inequalities of eignvalues and singular values is the main issue of the theory of matrices and perturbation analysis, which could be divivded into three main aspects: (1), the relations of the eignvalues and singular values of single matix; (2),the relations of eignvalues and singular values of several matrices; (3), the estimation of their bounds. Early, Wely, Courant, Fisher, Wielandt and Hoffaman have already done many works in these issues. In 1953, A.J.Hoffaman and H.W.Wielandt proposed the famous Wielandt-Hoffaman theorem, which described the relations of eignvalues of the difference of two symmatric matrices. Concretely, it associated the purturbation of eignvalues with the Enclid norm of the purturbation matrices, and it was the best conclusion for the error estimation which based on the float operation of the orthonormal transform.

Some interesting results have been obtained by[9],[12]. In the paper we strengthen H-W's and wu's results, considerably, we give their inverse inequality and their convenient proof, then we extend it to the inequality of singular values of arbitrary matrices.

Four sections are folded in this thesis. Section 1 gives us an introduction of the basic conception of H-W theorem and looks back the former works of this area. In section 2, we briefly introduce the classical H-W theorem and some other definitions. We discuss the symmetric form of H-W theorem and generalise it to the summation of matrix in section 3. The last part present the singular inequalities for the random matrices which include the difference and summation of the matrices and the numerical experiments are presented.

Key words: Singular values; Eigenvalues; Inequality.

目 录

中文摘要 i

英文摘要 ii

中文目录 iv

英文目录 v

第一节 引言 1

第二节 定义与预备知识 5

第三节 实对称矩阵和与差的特征值不等式 10

第四节 任意 $m \times n$ 阶矩阵和与差的奇异值不等式 22

参考文献 27

致 谢 30

Contents

Chinese Abstract i

English Abstract ii

Chinese Contents iv

English Contents v

Section I Introduction 1

Section II Some definitions and pre. knowledges 5

Section III The summation and difference of eigenvalue inequalities of real symmetry matrix 10

Section IV The summation and difference of eigenvalue inequalities of arbitrary matrix 22

References 27

Acknowledges 30

第一节 引 言

一　研究背景

矩阵有着悠久的发展历史和极其丰富的内容，作为一种基本的数学工具，矩阵在数学学科与其它科学技术领域，诸如数值分析、优化理论、微分方程、概率统计、运筹学、控制论、系统工程等学科都有广泛的应用，甚至在经济管理、社会科学等方面，矩阵也起着十分重要的作用. 现代科学技术的发展，特别是电子计算机技术的发展，为矩阵的应用开辟了更广阔的前景.

矩阵的奇异值和特征值的估计问题始终在数值代数和矩阵理论中占有十分重要的地位. 而矩阵最小奇异值和矩阵最大特征值的估计是矩阵分析中的重要课题之一，对其研究具有重要的理论意义和应用价值. 它们研究的问题一般包括：单个矩阵特征值 (奇异值) 之间的关系问题，多个矩阵之间特征值 (奇异值) 关系的问题，矩阵范数与特征值的关系问题，矩阵特征值 (奇异值) 的界的估计问题等等；例如在迭代求解线性方程组时，我们往往需要估计矩阵的谱条件数. 特别地，在线性方程组的解的扰动分析，特征值问题的扰动分析等问题中都会遇到矩阵的条件数. 因此给出矩阵条件数的估计对研究各种矩阵问题的扰动问题的扰动分析有重要意义，其中对奇异值的下界估计是一个估计矩阵谱条件数的关键的数. 奇异值的下界估计在其他许多领域中也是一个极重要的课题，因而有很重要的理论和实际应用价值，因此最小奇异值下界的估计一直是普遍关注的问题. 160 多年来，矩阵特征值 (奇异值) 问题的研究已经取得了许多丰硕的成果，文 [17] 、 [18] 、 [19] 、 [20] 、 [15] 、 [16] 等著作都对矩阵特征值问题的理论和方法进行了总结。

1964 年， A. S. Householder 在《 The theory of matrices in numerical analysis 》一文中作了详细的论述，主要讨论了一些基本工具，从矩阵分裂和模的理论的角度来分析特征值的局部化理论，并给出了值域的一般结果及若干不等式. 1965 年， J. H. Wilkinson 在《 The algebraic eigenvalue problem 》中作了系统而又精

辟的阐述，作为一本专门研究特征值的著作，作者用摄动理论和向后误差分析方法系统地论述代数特征值问题以及有关的线性代数方程组的各种解法，并对方法的性质用了透彻的分析. 接下来， Chatelin, G. H. Gloub ,C. F. Vanloan 等人又进一步完善了该理论.

矩阵特征值的计算与估计在理论和应用上都是十分重要的，但要精确计算特征值并非总有可能. 代数特征值问题一直以来都具有很重要的地位，虽然特征值问题具有貌似简单的提法，而且其基本理论多年来已为人们所熟知；然而欲求其精确解就会遇到各种挑战性问题。即使在某些特殊情况下有可能，付出的代价也会太大. 而在很多应用方面往往不必精确计算特征值，只需有一个粗略的估计就够了，或者给出它的一个上界或下界, 以达到我们预期的目的. 例如, 当研究一个迭代法的收敛性时, 便要判断迭代矩阵的特征值是否都落在单位圆内. 关于特征值界的估计, 已有一些经典的结论， 1902 年， Hirsch 给出了特征值的模、实部和虚部的范围， 1909 年， Schur 又提出了特征值估计的 Schur 不等式 [2], 还有众所周知的 Frobenius 估计式 [3] 等等. 这些估计式都是从矩阵的元素结构上给出了特征值的范围, 一般范围并不十分明确, 需要一些值的最大 (小) 值, 且估计的范围较大. 另外, 应用矩阵的迹来估计特征值的界，自 20 世纪 80 年代以来, 国内外已有大量的成果， [1] 、 [2] 、 [3] 、 [4] 是从矩阵特征值的均值的标准差与矩阵迹的关系来确定特征值的简单实用的上、下界； [5] 、 [6] 、 [7] 等是从矩阵的秩与迹的关系式来讨论问题的，由于前者没有后者那么多限制条件, 因而其结果的使用范围很广；在文献 [31] 、 [33] 中对非负阵的 Perron-roots 的上下界作了比较精确而又简单的估计等等.

以上研究的都是单个矩阵的特征值问题，而对于多个矩阵特征值、奇异值关系问题也有了大量的研究. 在早期， Wely, Courant 和 Fisher,Wielandt 和 Hoffaman 等为相关的研究做了大量的工作；另外文献 [32] 等也讨论了两个对称矩阵积的奇异值与特征值的上下界. 1953 年， A. J. Hoffaman 和 H. W. Wielandt 给出了两个矩阵差的特征值关系，即著名的 W-H 定理.

定理 1.1. 文献 [9] *W-H* 定理：

设 A,B,C, 都是 $n\times n$ 实对称阵，它们的特征值 (从大到小的次序排列) 分别为 $\alpha_1\geq\alpha_2\geq\cdots\geq\alpha_n$，$\beta_1\geq\beta_2\geq\cdots\geq\beta_n$，$\gamma_1\geq\gamma_2\geq\cdots\geq\gamma_3$. 若 B=C-A, 则

$$\sum_{i=1}^{n}(\gamma_i-\alpha_i)^2\leq\sum_{i=1}^{n}\beta_i^2=\|B\|_F^2. \tag{1.1}$$

A. J. Hoffaman 和 H. W. Wielandt 是基于线性规划的理论证明了这一定理，后来 J. H. Wilkinson 在《 The algebraic eigenvalue problem 》一文中用纯代数的方法证明了这一定理，该方法虽然不甚巧妙但较为初等，它实际上应该归功于 Givens(1954); 蒋尔雄在文 [10] 中给出了 W-H 定理的一个简化证明，孙继广在《矩阵扰动分析》也给出了另一种证明方法.

这一定理很快的成为数值代数和计算数学的中心内容之一，并引起许多学者的兴趣，它有着比较广泛的应用，特别是在误差估计方面．设 A 是一个方阵， A+B 是它的扰动矩阵， λ 是 A 的特征值，$\overline{\lambda}$ 是 A+B 的特征值，关于特征值的传统误差界是估计 $|\lambda-\overline{\lambda}|$ 的上界，对此上界的估计有多种类型，而其中一种就是 Wielandt-Hoffaman 型的，具体的如下：

1)Bauer-Fike 型：用谱范数界定矩阵特征值与其扰动矩阵特征值的误差界．即 $\min|\lambda-\overline{\lambda}|$

2)Wielandt-Hoffaman 型：用 Frobenius 范数界定矩阵特征值与对应的按一定顺序排列的其扰动矩阵特征值的所有距离的平方的平方根．即：

$$min\sqrt{\sum_{i=1}^{n}|\lambda-\overline{\lambda}|^2} \tag{1.2}$$

也就是若记 C=A+B, B 的特征值为 γ_i. 则有

$$\sum_{i=1}^{n}(\lambda-\overline{\lambda})^2\leq\sum_{i=1}^{n}\gamma_i^2=\|B\|_F^2. \tag{1.3}$$

3)Weyl 型：用谱范数界定所有的矩阵与其扰动矩阵的第 i 个特征值之间距离的最大者，即 $\max|\lambda-\overline{\lambda}|$, 其中矩阵的特征值按非减的顺序排列.

以上的测量误差的方法统称为绝对扰动界,绝对扰动界可以转化为相对扰动界,相对扰动界通常用以下三种方式来度量：$\frac{|\lambda-\bar{\lambda}|}{|\lambda|}$，$\frac{|\lambda-\bar{\lambda}|}{\sqrt{|\lambda\bar{\lambda}|}}$，$\frac{|\lambda-\bar{\lambda}|}{\sqrt[p]{|\lambda|^p+|\bar{\lambda}|^p}}$. 另外在 [21] 中，孙继广研究了正规矩阵的谱扰动，给出了一个 Hoffman-Wielandt 型不等式， [22] 将 [21] 中结果加以推广，得到了可对角化矩阵的相应扰动定理. 后来，这方面的研究工作又取得了一些新的成果 [23] 、 [24] 。在文 [25] 中，建立几个矩阵范数不等式，然后将它们用于可对角化矩阵 (正规矩阵) 的谱扰动，导出几个新的 H-W 型扰动. 因此这个定理在矩阵扰动方面有着广泛的应用，文 [27] 孙继广关于 Wielandt-Hoffaman 定理和文 [26] 伍俊良关于 Wielandt-Hoffaman 定理的一些结果等也推广了 W-H 定理的结果.

本文所作的研究主要是在 W-H 定理的基础上，找出它的对称形式，并推广到了奇异值的情形. 首先，在 H-W 定理及文 [12] 的基础上，更进一步的得到了 H-W 的反向不等式, 即它的对称形式, 并结合控制论的知识给出了相当简洁的证明, 同时把文 [12] 中矩阵和的特征值不等式推广，同样得到了它的对称形式，即它的下界. 最后，把原先讨论的对称阵的特征值的不等式推广到任意的 $m\times n$ 阶矩阵的奇异值不等式，在原有的不等式上利用构造对称阵的方法巧妙的证明了奇异值不等式的成立，并给出了数值例子.

第二节　定义及预备知识

在本节开始之前，我们先引入以下几个常用也是最基本的概念，然后给出了几个必要的引理和定理.

定义 2.1. $A=(a_{ij})$ 是实对称的，如果 $a_{ij}\in R$, 且 $a_{ij}=a_{ji}, i=1,\cdots,n; j=1,\cdots,n$.

定义 2.2. *Hermite* 矩阵 A^*A 的特征值的非负平方根称为矩阵 $A\in C^{m\times n}$ 的奇异值.

定义 2.3. 设矩阵 $A\in C^{n\times n}$, 使得 $Ax=\lambda x$, 则 λ 叫 A 的特征值， x 叫做 A 的属于特征值 λ 的特征向量. A 的所有特征值的全体，叫做 A 的谱 *(Spectrum)*, 记作 λ_A.

定义 2.4. 对于 $A=(a_{ij})\in C^{m\times n}$, 令 $\|A\|_F=\sqrt{\sum_{i,j}|a_{ij}|^2}$, 则它是一种范数，称作矩阵 A 的 *Frobenius* 范数.

定义 2.5. 对任意的 $\alpha=(\alpha_1,\alpha_2,\alpha_3,\cdots,\alpha_n)^\top\in R^n$, 记它按从大到小的顺序排列为 $\alpha^{[1]}\geq\alpha^{[2]}\geq\cdots\alpha^{[n]}$. 设 $\alpha,\beta\in R^n$, 且向量 β 中的元素按从大到小的顺序排列为 $\beta^{[1]}\geq\beta^{[2]}\geq\cdots\beta^{[n]}$, 若满足

$$\sum_{i=1}^{k}\alpha^{[i]}\leq\sum_{i=1}^{k}\beta^{[i]},\qquad k=1,2,\cdots,n. \tag{2.1}$$

则称 β 弱控制 α, 记为 $\alpha\preceq\beta$. 若 $\alpha\preceq\beta$, 且

$$\sum_{i=1}^{n}\alpha^{[i]}=\sum_{i=1}^{n}\beta^{[i]}. \tag{2.2}$$

则称 β 控制 α, 记为 $\alpha\prec\beta$.

在引入了上述的几个定义之后，我们来看 Hoffaman-Wielandt 定理

定理 2.1. ($H-W$ 定理)

设 B=C-A, A,B,C, 都是 $n\times n$ 实对称阵，它们的特征值 (从大到小的次序排列) 分别为 $\alpha_i,\beta_i,\gamma_i$ $(i=1,2,\cdots,n)$. 则 A,B,C 的特征值之间有如下的关系成立：

$$\sum_{i=1}^{n}(\gamma_i-\alpha_i)^2\leq\sum_{i=1}^{n}\beta_i^2=\|B\|_F^2. \tag{2.3}$$

证明： 可参见文献 [9]; [10]; [15], 另外在文献 [35] 中给出了它的一种简化证明, 其证法将在定理 3.1, 3.2 中得以体现. □

另外在文献 [12] 伍俊良，刘飞. 实对称矩阵和与差的一些特征值与 F- 范数不等式一文中，给出了两个矩阵和的特征值不等式.

定理 2.2. (参见文献 [12])

设 A,B,C, 都是 $n\times n$ 实对称阵，它们的特征值 (从大到小的次序排列) 分别为 $\alpha_1\geq\alpha_2\geq\cdots\geq\alpha_n$, $\beta_1\geq\beta_2\geq\cdots\geq\beta_n$, $\gamma_1\geq\gamma_2\geq\cdots\geq\gamma_3$. 如果 B=C+A, 则 A,B,C 的特征值之间有如下的关系成立：

$$\sum_{i=1}^{n}\beta_i^2\leq\sum_{i=1}^{n}(\gamma_i+\alpha_i)^2. \tag{2.4}$$

证明： 对于本定理可分为两种情况加以证明，即 A 为正定的特殊情况和 A 为非正定的情况.

1) 假设 A 是正定的, 由于 A ， B ， C 都是 $n\times n$ 实对称阵, 则分别存在正交阵 P_1,P_2,P_3, 使得： $A=P_1^\top diag[\alpha_1,\alpha_2\cdots\alpha_n]P_1$; $B=P_2^\top diag[\beta_1,\beta_2\cdots\beta_n]P_2$; $C=P_3^\top diag[\gamma_1,\gamma_2\cdots\gamma_n]P_3$. 由于 A 是正定的，则有 $\alpha_i\geq 0$, $(i=1,2,\cdots,n)$. 显然有

$$TrB^2=Tr(P_2^\top diag[\beta_1^2,\beta_2^2\cdots\beta_n^2]P_2)=\sum_{i=1}^{n}\beta_i^2. \tag{2.5}$$

同理

$$TrA^2 = \sum_{i=1}^{n} \alpha_i^2 \qquad TrC^2 = \sum_{i=1}^{n} \gamma_i^2 \tag{2.6}$$

又

$$TrB^2 = Tr(C+A)^2 = Tr(C^2 + A^2 + 2AC) = \sum_{i=1}^{n} \gamma_i^2 + \sum_{i=1}^{n} \alpha_i^2 + 2\sum_{i=1}^{n}(AC)$$

从而欲证 (2.4) 式成立，只需证明 $Tr(AC) \le \sum\limits_{i=1}^{n}(\alpha_i\beta_i)$ 即可，而

$$\begin{aligned} tr(AC) &= tr(P_1^{\top} diag(\alpha_1, \alpha_2, \cdots, \alpha_n)]P_1P_3^{\top} diag[\gamma_1, \gamma_2 \cdots \gamma_n]P_3) \\ &= tr(diag(\alpha_1, \alpha_2, \cdots, \alpha_n)]P_1P_3^{\top} diag[\gamma_1, \gamma_2 \cdots \gamma_n]P_3P_1^{\top}) \end{aligned}$$

记 $P_1P_3^{\top} = P = p(ij)$, 易知 P 仍为正交阵，则

$$\sum_{i=1}^{n} p(ij) = 1 \ (j = 1, 2, \cdots, n), \qquad \sum_{j=1}^{n} p(ij) = 1 \ (i = 1, 2, \cdots, n). \tag{2.7}$$

故有

$$\begin{aligned} tr(AC) &= tr(diag(\alpha_1, \alpha_2, \cdots, \alpha_n)]P diag[\gamma_1, \gamma_2 \cdots \gamma_n]P^{\top}) \\ &= (\alpha_1, \alpha_2 \cdots \alpha_n) \begin{pmatrix} p_{11}^2 & p_{12}^2 & \cdots & p_{1n}^2 \\ p_{21}^2 & p_{22}^2 & \cdots & p_{2n}^2 \\ \cdots & \cdots & \cdots & \cdots \\ p_{n1}^2 & p_{n2}^2 & \cdots & p_{nn}^2 \end{pmatrix} \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_n \end{pmatrix} \end{aligned}$$

令

$$\begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{pmatrix} = \begin{pmatrix} p_{11}^2 & p_{12}^2 & \cdots & p_{1n}^2 \\ p_{21}^2 & p_{22}^2 & \cdots & p_{2n}^2 \\ \cdots & \cdots & \cdots & \cdots \\ p_{n1}^2 & p_{n2}^2 & \cdots & p_{nn}^2 \end{pmatrix} \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_n \end{pmatrix}$$

易知

$$\sum_{i=1}^{n} \xi_i = \sum_{i=1}^{n} \gamma_i. \tag{2.8}$$

同时对任意的 $k,(1\le k\le n)$ 有

$$\begin{aligned}\sum_{i=1}^{k}\xi_i=\sum_{i=1}^{k}\sum_{j=1}^{n}p_{ij}^2\gamma_j &= \sum_{i=1}^{k}\gamma_i-\sum_{i=1}^{k}(1-\sum_{j=1}^{n}p_{ij}^2)\gamma_i+\sum_{i=1}^{k}\sum_{j=i+1}^{n}p_{ij}^2\gamma_j\\ &\le \sum_{i=1}^{k}\gamma_i-\gamma_k\sum_{i=1}^{k}(1-\sum_{j=1}^{n}p_{ij}^2)+\gamma_k\sum_{i=1}^{k}\sum_{j=i+1}^{n}p_{ij}^2\\ &= \sum_{i=1}^{k}\gamma_i-\gamma_k\sum_{i=1}^{k}(1-\sum_{j=1}^{n}p_{ij}^2)+\gamma_k(1-\sum_{j=1}^{k}p_{ij}^2)\\ &= \sum_{i=1}^{k}\gamma_i.\end{aligned} \tag{2.9}$$

故对任意的整数 k 有 $\sum_{i=1}^{k}\xi_i\le\sum_{i=1}^{k}\gamma_i,\ k=1,2,\cdots n.$ 成立，而 $\alpha_i>0$ 为正数，所以有

$$Tr(AC)=\sum_{i=1}^{n}(\alpha_i\xi_i)\le\sum_{i=1}^{n}\alpha_i\gamma_i \tag{2.10}$$

成立，从而

$$\sum_{i=1}^{n}\beta_i^2\le\sum_{i=1}^{n}(\gamma_i+\alpha_i)^2.$$

2) 如果 A 非正定，则存在充分大的实数 $d>0$, 使得 $A+dI$ 为正定阵 (I 为 $n\times n$ 阶的单位阵， $B+dI$ 可为正定阵也可为非正定阵), 则由 1) 有

$$\sum_{i=1}^{n}(\beta_i+d)^2\le\sum_{i=1}^{n}(\gamma_i+\alpha_i+d)^2=\sum_{i=1}^{n}[(\gamma_i+\alpha_i)+d]^2. \tag{2.11}$$

将 $\sum_{i=1}^{n}\beta_i=TrB=Tr(C+A)=\sum_{i=1}^{n}(\gamma_i+\alpha_i)^2$ 代入 (2.11), 有

$$\sum_{i=1}^{n}\beta_i^2\le\sum_{i=1}^{n}(\gamma_i+\alpha_i).$$

□

故结论成立.

本定理的证明过程可参见文献 [12](伍俊良，刘飞. 实对称矩阵和与差的一些特征值与 F- 范数不等式). 不难看出这一结果与 Wielandt- Hoffaman 的结果具有某